

January 13, 2025

Hon. Ona T. Wang  
United States Magistrate Judge  
Southern District of New York

Re: *The New York Times Company v. Microsoft Corp.* No.: 23-cv-11195; *Daily News, LP v. Microsoft Corp.*: No. 1:24-cv-3285; *The Center for Investigative Reporting, Inc. v. OpenAI, Inc.*: No. 1:24-cv-04872: Dispute Regarding Microsoft’s Custodial Search Terms

Dear Magistrate Judge Wang:

Plaintiffs The New York Times Company (“The Times”), Daily News, LP et al. (“Daily News”), and the Center for Investigative Reporting (“CIR”) (collectively, the “News Plaintiffs”) request that this Court address two ripe disputes concerning Microsoft’s collection and review of custodial documents.

Specifically, the News Plaintiffs request an order (1) setting a target search term hit count cap of 600,000 custodial documents for Microsoft to collect and review, and (2) requiring Microsoft, as part of that custodial review process, to collect documents dated through July 1, 2024.<sup>1</sup> Resolution of these disputes is necessary for discovery to proceed in a timely fashion.

### **1. The Court Should Set a Target Hit Count Cap of 600,000 Custodial Documents.**

After over six months of negotiating search term proposals, it is time for these negotiations to conclude so the parties can move on to the next phase of the case.

The best path forward is a familiar one: this Court should resolve this dispute the same way it resolved the News’ Plaintiffs’ analogous dispute with OpenAI—that is, to impose a target hit count cap and direct the News Plaintiffs to tailor their proposals to meet that target. *See* Dkt. 304 at 3 (imposing a target hit count cap of 500,000 additional documents for OpenAI to review); Dkt. 365 at 4 (increasing target hit count to 600,000 after OpenAI was ordered to collect documents from eight additional custodians).<sup>2</sup> This approach will allow the News Plaintiffs and Microsoft to

---

<sup>1</sup> The parties have met and conferred about search terms on multiple occasions by email and videoconference, including most recently on January 9, 2025. *See also* Ex. 1 (emails between counsel). During that meet-and-confer, the parties discussed the two disputes summarized in this brief, but were unable to reach agreement.

<sup>2</sup> The News Plaintiffs and OpenAI currently dispute “whether the 600,000 hit count ‘cap’ applies to all custodial documents produced to date or only those produced after the October 30, 2024, conference.” Dkt. 365 at 4; *see also* Dec. 3, 2024 Tr. at 61:3-64:21. Under the News Plaintiffs’ interpretation of the Court’s October 31, 2024 Order, the total hit count cap for OpenAI would fall at around 730,000, since OpenAI had reviewed approximately 130,000 documents before The Times filed the motion to compel which resulted in the order compelling OpenAI to review 600,000 additional documents. *See* Dkt. 228-1 at 1; *see also* Dkt. 233-2 (OpenAI declarant characterizing The Times’s search term motion as seeking OpenAI’s review of “new documents (i.e., not previously reviewed by OpenAI)”). The News Plaintiffs’ most recent search term proposal yielded approximately 667,000 documents, which essentially splits the difference between OpenAI’s understanding of the Court’s Order (600,000) and the News Plaintiffs’ understanding

efficiently complete this process without wasting time disputing individual terms. This Court's previous selection of a target cap has already enabled the News Plaintiffs and OpenAI to make significant progress toward concluding search term negotiations. *See* Dec. 3, 2024 Hearing Tr. at 60:23-25 (Counsel for News Plaintiffs noting that "We have been exchanging search term proposals [with OpenAI], and we've seen a lot of progress with the parties agreeing on terms because of the hit count cap"); *id.* at 60:14-15 (Counsel for OpenAI noting that "So we should be able to get to 600 continuing to sort of slim down to the target weight"). Microsoft should be subject to the same approach, which promises the same productive results.

The target of 600,000 custodial documents is proportional given Microsoft's substantial role in these cases. As explained in The Times's Complaint, Microsoft has been, and continues to be, intimately involved in the training, development, and commercialization of Defendants' large-language models. FAC ¶ 66. Relatedly Microsoft has developed and commercialized user-facing products, including Copilot, which violate the News Plaintiffs' copyrights through the substitutive outputs they provide in response to user prompts. FAC ¶¶ 112-17. Evidence of this and other conduct is relevant to the claims against Microsoft for direct, contributory, and vicarious copyright infringement, DMCA violations, trademark infringement, and unjust enrichment. *E.g.*, Ex. 2 at 65 (July 2023 email exchange discussing [REDACTED]).

The hit count target is also supported by caselaw. *See, e.g., King v. Habib Bank Ltd.*, 2023 WL 3558773, at \*2 (S.D.N.Y. Apr. 21, 2023) (Wang, M.J.) (ordering the defendant to review additional documents, notwithstanding the defendant had already "review[ed] approximately 750,000 documents for prior search terms"); *Fort Worth Employees' Ret. Fund v. J.P. Morgan Chase & Co.*, 297 F.R.D. 99, 103 (S.D.N.Y. 2013) (ordering the defendant to apply additional search terms and review additional documents, notwithstanding the defendant "ha[d] already conducted a search for ESI that yielded some 875,000 document hits").

## **2. Microsoft Should Collect and Review Documents Dated Through July 1, 2024.**

The News Plaintiffs and Microsoft also dispute the date range that Microsoft has applied to its custodial document collections. Microsoft has limited its collections to documents dated on or before December 27, 2023, which is the same day The Times filed its case. Put differently, and subject to limited exceptions,<sup>3</sup> Microsoft is using the filing date of the News Plaintiffs' first-filed complaint as the cutoff for custodial document productions for all three News cases.

Microsoft's date cutoff is improper. Documents that post-date The Times's Complaint are relevant, including because The Times alleges ongoing misconduct. *See* FAC ¶ 167 (describing "Defendants' ongoing infringing conduct"); *id.* ¶ 66 (alleging that Microsoft "continues to"

---

(730,000). The News Plaintiffs have therefore proposed that OpenAI accept this most recent proposal, and that parties be done proposing additional search terms, absent good cause.

<sup>3</sup> Microsoft has stated that, in some cases, "where appropriate and relevant," it has opted to review documents dated through June 2024. Ex. 1 at 13. This vague and ad hoc approach will invariably lead to many more disputes, as the parties continue to disagree about what is "relevant." The more efficient and equitable path is to impose an across-the-board cutoff date for document review, including because the cutoff date is tied to a target hit count cap, such that any extension of the cutoff date will not result in Microsoft reviewing more documents than it otherwise would have.

collaborate with OpenAI to infringe Times works). “Rule 26 does not limit discovery to pre-complaint events in cases such as this in which the complaint alleges that [the wrongful conduct] was continuing as of the date the complaint was filed.” *Robco Distributors, Inc. v. Gen. Glass Int’l Corp.*, 101 F.R.D. 547, 548 (E.D.N.Y. 1984); *see also Deluxe Fin. Servs., LLC v. Shaw*, 2017 WL 7369890, at \*5 (D. Minn. Feb. 13, 2017) (“it may be appropriate to allow post-complaint discovery where documents requested are relevant to the claims and defenses in the case, especially where the complaint alleges continuing violations”).

In any event, even if it were proper to use the Complaint’s filing date, Microsoft completely ignores the *Daily News* case, filed on April 30, 2024, and the *CIR* case, filed on June 27, 2024. Since consolidation, the News Plaintiffs have collaborated to jointly propose search terms and custodians to Defendants, and these search terms are being used to collect and review Defendants’ documents for production across all three cases. This approach has been productive in streamlining discovery. However, applying a cutoff date that *predates* two of the three consolidated lawsuits necessarily risks prejudicing the Daily News and CIR plaintiffs. Therefore, the cutoff should be July 1, 2024, which is right after the final News case was filed (*CIR*), on June 27, 2024.

Finally, Microsoft’s objection to this cutoff is particularly misplaced because this cutoff will not result in any additional burden on Microsoft to review additional documents. The News Plaintiffs’ proposed 600,000 cap would include any documents based on the July 2024 cutoff and will count toward the 600,000 target cap.

Respectfully,

/s/ Ian Crosby

Ian B. Crosby  
Susman Godfrey L.L.P.

/s/ Steven Lieberman

Steven Lieberman  
Rothwell, Figg, Ernst & Manbeck

/s/ Matt Topic

Matt Topic  
Loevy & Loevy

cc: All Counsel of Record (via ECF)